Online appendix for Attendance Boundary Policies and the Limits to Combating School Segregation

A Construction of SES index

This appendix outlines how we construct our socio-economic index. We describe our approach of reducing a set of socio-economic variables to a single socio-economic index (SES index henceforth) and we evaluate the index' performance.

We construct our SES-index by choosing the first variable resulting from a principal component analysis (PCA) based on the following variables:

- *INC*: We calculate the market income rank of all adults in the population. We select the highest income rank observed in a household.
- *LCE*: A dummy which takes the value of one if an adult in a household has completed a long cycle education.
- *NE*: A dummy which takes the value of one if an adult in a household has not completed in education beyond primary school or have no registered education.
- *EMP*: A dummy which takes the value of one if an adult in a household is employed.

We select the first component of the PCA. This leads to the following index:

$$SES = 0.62INC + 0.38LCE - 0.44NE + 0.53EMP, \tag{1}$$

where all variables have been standardized to their corresponding z-scores. This index accounts for 47 pct. of the variation in the four variables. The SES-index applied in our paper is the population ranks of SES, as such it is uniformly distributed on the unit interval.

To get a sense of the mapping between the underlying variables we calculate averages of the underlying variables in percentiles of the SES-index. The results are displayed in Figure A.1. While this is a very simple index we find that this component is intuitive. In the bottom of the distribution almost all households have an uneducated parent and no parent with a high cycle education. In the top 75 percent of the distribution no household contain an uneducated parent. Income and employment are both rising in the SES-index. Thus we find it safe to assume that the SES-index reflects a true underlying socioeconomic status.

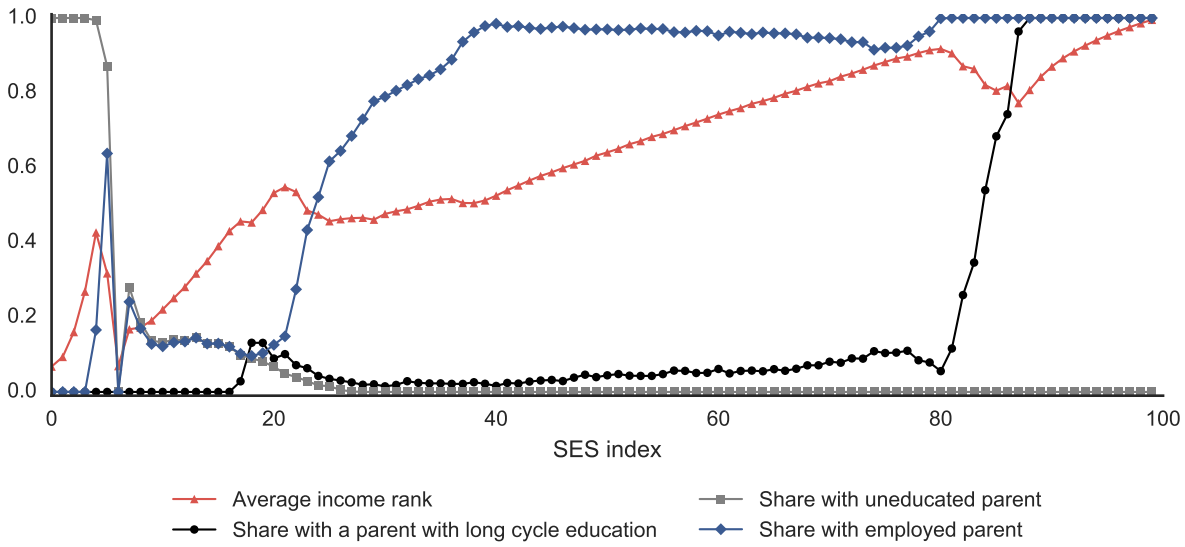


Figure A.1: Average characteristics as a function of SES-index

The figure depicts means of variables used to construct the SES-index. The SES-index is uniformly distributed on the unit interval. Each marker represent the mean of the variable in question within a percentile bin. Income rank is bounded between 0 and 1.

B Supplementary tables and plots

Table B.1: Treatment propensity and characteristics for municipalities and schools

	(1) Municipal level	(2) Attendance zone level
SES (mean)	0.041	0.092
	(0.346)	(0.161)
SES (std. dev.)	0.712	0.949
	(1.731)	(0.551)
Population density (mean)	0.091	0.037
	(0.029)	(0.010)
Population density (std. dev.)	-0.121	0.098
	(0.090)	(0.041)
N	88	1083
F.E.	-	Municipal
Clusters	-	Municipal

This table contains models where the dependent variables measure the extent of change in attendance boundaries. The indicator variable is computed at the attendance boundary level and measures whether the attendance zone has experienced changes to its boundaries over the period 2008-2015. For municipalities, the outcome is the average of the attendance-zone indicator variables. The explanatory variables are descriptive statistics computed at the municipalities and attendance-zone level.

Table B.2: Alternative measure of enrollment in assigned and changes to school characteristics - far away matching

	(1)	(2)	(3)	(4)
	– Subsample: Unmatched sample –			
T×Post	-0.278	-0.281	-0.289	-0.283
	(0.032)	(0.035)	(0.035)	(0.037)
ΔDistance×Post	-0.080	-0.079	-0.116	-0.098
	(0.029)	(0.033)	(0.085)	(0.062)
ΔSchool SES×Post	1.458			2.254
	(0.217)			(0.669)
ΔNon-Western share×Post		-0.828		0.522
		(0.151)		(0.402)
ΔSchool Value-added×Post			-0.047	-0.014
			(0.051)	(0.054)
N	48,458	48,458	47,048	47,048
	– Subsample: Matched sample –			
T×Post	-0.281	-0.285	-0.291	-0.283
	(0.037)	(0.040)	(0.033)	(0.034)
ΔDistance×Post	-0.065	-0.062	-0.134	-0.114
	(0.023)	(0.024)	(0.084)	(0.063)
ΔSchool SES×Post	1.260			1.952
	(0.214)			(0.706)
ΔNon-Western share×Post		-0.758		0.463
		(0.166)		(0.432)
ΔSchool Value-added×Post			-0.099	-0.069
			(0.060)	(0.059)
N	141,681	141,681	138,018	138,018

This table shows the coefficient of the model in Equation (4) where the alternative compliance measure, enrollment in assigned school at age 7, is the dependent variable. See Subsection 5.4 under the heading "Comparing compliance for treatment and control groups" for a discussion of outcome variable definition. It uses both the full/unmatched and the matched sample where children outside of the district are used as controls. Details of the matching procedure are found in Subsection 5.4 under the heading "Spillovers to untreated". Columns 2-4 display regression results for the model presented in Equation (4) for one school characteristic at a time. Column 5 displays the result of an estimation using all characteristics at a time. The models are estimated with origin-attendance-zone by year fixed effects and origin-attendance-zone-by-new-attendance-zone fixed effects. Standard errors are in parentheses and clustered by the actual municipality for treated units and for control units the municipality that they are matched to.

Table B.3: Enrollment in the assigned school and change in school characteristics - full model specification

	(1)	(2)	(3)	(4)
Distance to closest school	0.003	0.003	0.003	0.003
	(0.002)	(0.002)	(0.002)	(0.002)
ΔDist	0.016	0.014	0.049	0.033
	(0.009)	(0.008)	(0.029)	(0.019)
ΔSES	-0.582			-0.762
	(0.336)			(0.589)
ΔNW		0.524		-0.048
		(0.159)		(0.356)
ΔSVA			0.108	0.084
			(0.048)	(0.048)
T×Post	0.350	0.345	0.330	0.344
	(0.023)	(0.027)	(0.020)	(0.021)
ΔDist×Post	-0.054	-0.052	-0.090	-0.062
	(0.023)	(0.023)	(0.070)	(0.054)
ΔSES×Post	0.888			0.945
	(0.182)			(0.339)
ΔNW×Post		-0.543		0.005
		(0.156)		(0.229)
ΔSVA×Post			0.040	0.075
			(0.042)	(0.046)
N	43,450	43,450	42,041	42,041

This table contains the full estimation output of Table 3. Each column in this table displays regression results for the model presented in Equation (4). The dependent variable for these models is compliance, i.e, enrollment in the assigned school where control units are artificially assigned nearby recipient schools. Column 1 omits all measures of change in school characteristics, only change in distance from the original to the newly assigned school. Columns 2-4 display uses one of three school characteristics at a time. Column 5 displays the result of an estimation using all characteristics at a time. Standard errors are in parentheses and clustered on the municipality level.

Table B.4: Responses along margins for non-enrollment in the assigned school - full model specification

	Comply	Original	Public	Private	Move
Distance to closest school	0.004	-0.046	0.021	0.024	-0.002
	(0.002)	(0.007)	(0.005)	(0.008)	(0.001)
ΔDistance	0.013	0.025	-0.030	0.004	-0.012
	(0.009)	(0.025)	(0.017)	(0.011)	(0.005)
ΔSchool SES	-0.374	-1.411	1.062	1.631	-0.908
	(0.553)	(0.507)	(0.378)	(0.433)	(0.565)
T×Post	0.367	-0.143	-0.221	-0.021	0.018
	(0.024)	(0.032)	(0.016)	(0.008)	(0.010)
ΔDistance×Post	-0.062	0.024	0.043	-0.016	0.010
	(0.024)	(0.022)	(0.021)	(0.013)	(0.006)
ΔSchool SES×Post	0.924	0.216	-0.888	-0.237	-0.015
	(0.190)	(0.157)	(0.144)	(0.091)	(0.066)
N	43,393	43,393	43,393	43,393	43,393

This table contains the full estimation output of panel a) of Table 4. Columns display regression results for the model presented in Equation (4) using SES as a measure of schools' characteristics. In column 1 the dependent variable is compliance, i.e, enrollment in the assigned school where control units are artificially assigned nearby recipient schools. The other columns capture one of the margins of non-compliance from Equation (2), see the columns' titles. The models are estimated with homogeneous compliance both in the sign of change and household-SES. Moreover, the models use origin-attendance-zone by year fixed effects and origin-attendance-zone-by-new attendance-zone fixed effects. Standard errors are in parentheses and clustered on the level of origin-attendance-zone. The remaining output is available upon request.

Table B.5: Responses along margins for non-enrollment in the assigned school - fully interacted with household SES

	Comply	Original	Public	Private	Move
T×Post× Household SES Q1	0.459	-0.184	-0.294	-0.009	0.028
	(0.066)	(0.077)	(0.029)	(0.022)	(0.025)
T×Post×Household SES Q2	0.397	-0.128	-0.220	-0.053	0.004
	(0.039)	(0.043)	(0.037)	(0.034)	(0.012)
T×Post×Household SES Q3	0.336	-0.103	-0.228	-0.022	0.017
	(0.034)	(0.067)	(0.035)	(0.061)	(0.012)
T×Post×Household SES, Q4	0.354	-0.098	-0.241	-0.034	0.018
	(0.027)	(0.030)	(0.030)	(0.032)	(0.012)
ΔDistance×Post×Household SES Q1	-0.063	0.050	0.041	-0.046	0.018
	(0.026)	(0.037)	(0.027)	(0.020)	(0.011)
ΔDistance×Post×Household SES Q2	-0.058	0.017	0.053	-0.022	0.010
	(0.031)	(0.029)	(0.029)	(0.019)	(0.005)
ΔDistance×Post×Household SES Q3	-0.062	0.024	0.041	-0.013	0.011
	(0.023)	(0.020)	(0.019)	(0.016)	(0.008)
ΔDistance×Post×Household SES, Q4	-0.080	0.015	0.037	0.022	0.006
	(0.028)	(0.024)	(0.027)	(0.020)	(0.007)
Δ^-School SES×Post×Household SES, Q1	1.368	0.174	-1.575	0.016	0.018
	(0.518)	(0.559)	(0.390)	(0.457)	(0.146)
Δ^-School SES×Post×Household SES Q2	0.807	0.071	-0.691	-0.066	-0.121
	(0.240)	(0.324)	(0.256)	(0.311)	(0.063)
Δ^-School SES×Post×Household SES Q3	0.749	0.775	-1.187	-0.374	0.037
	(0.241)	(0.479)	(0.365)	(0.555)	(0.065)
Δ^-School SES×Post×Household SES Q4	1.176	0.504	-1.060	-0.597	-0.024
	(0.184)	(0.304)	(0.247)	(0.170)	(0.040)
Δ^+School SES×Post×Household SES, Q1	0.036	0.532	-0.194	-0.258	-0.115
	(0.213)	(0.291)	(0.299)	(0.241)	(0.171)
Δ^+School SES×Post×Household SES Q2	1.517	-0.719	-0.876	0.123	-0.046
	(0.439)	(0.281)	(0.360)	(0.323)	(0.198)
Δ^+School SES×Post×Household SES Q3	1.072	-0.153	-1.036	0.000	0.117
	(0.422)	(0.323)	(0.419)	(0.614)	(0.108)
Δ^+School SES×Post×Household SES Q4	0.867	0.083	-0.704	-0.336	0.090
	(0.584)	(0.254)	(0.621)	(0.197)	(0.183)
N	43,393	43,393	43,393	43,393	43,393

This table expands the model of panel c) in Table 4 with household-SES interactions with both the response to changes in distance and the overall compliance rate after redrawn boundaries. Columns display regression results for the model presented in Equation (4) using SES as a measure of schools' characteristics. In column 1 the dependent variable is compliance, i.e, enrollment in the assigned school where control units are artificially assigned nearby recipient schools. The other columns capture one of the margins of non-compliance from Equation (2), see the columns' titles. The models are estimated with homogeneous compliance both in the sign of change and household-SES. Moreover, the models use origin-attendance-zone by year fixed effects and origin-attendance-zone-by-new-attendance-zone fixed effects. Standard errors are in parentheses and clustered on the level of origin-attendance-zone. Remaining output is available upon request.

Table B.6: Responses along different enrollment margins from changes to non-Western-share, interacted with own ethnicity

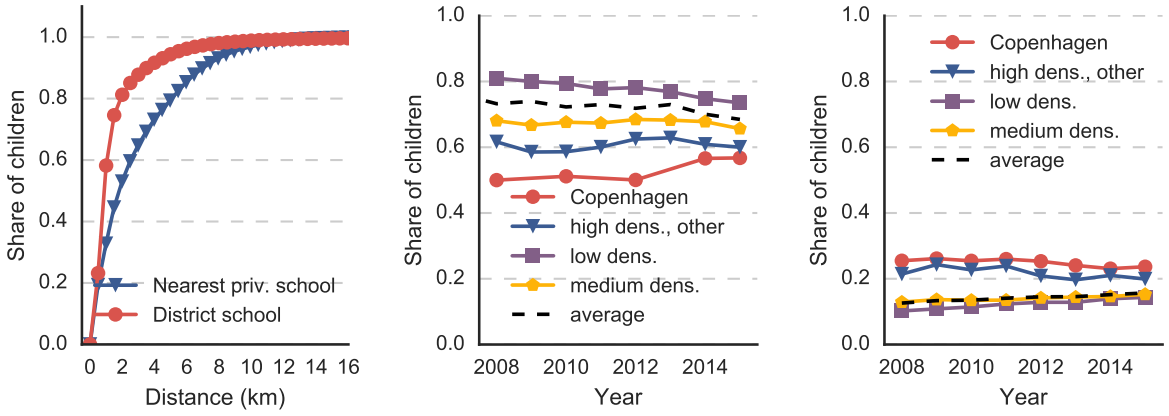
	Comply	Original	Public	Private	Move
T×Post×Western household	0.371	-0.148	-0.223	-0.020	0.020
	(0.026)	(0.027)	(0.018)	(0.011)	(0.011)
ΔDistance×Post×Western household	-0.059	0.024	0.039	-0.014	0.010
	(0.023)	(0.021)	(0.020)	(0.013)	(0.006)
ΔWestern share×Post×Western household	-0.704	-0.185	0.776	0.093	0.020
	(0.175)	(0.111)	(0.147)	(0.072)	(0.065)
T×Post×Non-Western household	0.355	-0.131	-0.219	-0.004	-0.000
	(0.040)	(0.037)	(0.025)	(0.030)	(0.011)
ΔDistance×Post×Non-Western household	-0.076	0.102	-0.005	-0.031	0.010
	(0.032)	(0.101)	(0.074)	(0.084)	(0.008)
ΔNon-Western share×Post×Non-Western household	-0.281	-0.317	0.484	0.049	0.065
	(0.189)	(0.226)	(0.110)	(0.084)	(0.081)
N	43,393	43,393	43,393	43,393	43,393

Columns display regression results for the model presented in Equation (4) using the share of non-Western children as a measure of schools' characteristics. In column 1 the dependent variable is compliance, i.e, enrollment in the assigned school where control units are artificially assigned nearby recipient schools. The other columns capture one of the margins of non-compliance from Equation (2), see the columns' titles. Characteristics are interacted with a household non-Western dummy. The models are estimated with origin-attendance-zone by year fixed effects and origin-attendance-zone-by-new-attendance-zone fixed effects. Standard errors are in parentheses and clustered on the level of origin-attendance-zones.

Table B.7: Compliance as a function of change in education and income

	Comply	Original	Public	Private	Move
	– a) Changes in share with long-cycle education –				
T×Post	0.357	-0.148	-0.209	-0.017	0.016
	(0.019)	(0.031)	(0.017)	(0.008)	(0.010)
ΔDistance×Post	-0.061	0.024	0.044	-0.017	0.010
	(0.024)	(0.024)	(0.023)	(0.013)	(0.006)
ΔEducation share×Post	1.347	0.254	-1.074	-0.479	-0.048
	(0.252)	(0.200)	(0.225)	(0.107)	(0.094)
N	43,393	43,393	43,393	43,393	43,393
	– b) Changes in average income –				
T×Post	0.368	-0.143	-0.222	-0.020	0.018
	(0.024)	(0.031)	(0.016)	(0.008)	(0.010)
ΔDistance×Post	-0.062	0.024	0.042	-0.016	0.011
	(0.024)	(0.022)	(0.021)	(0.013)	(0.006)
ΔAverage income×Post	1.124	0.249	-1.103	-0.253	-0.018
	(0.247)	(0.203)	(0.194)	(0.113)	(0.086)
N	43,393	43,393	43,393	43,393	43,393

Columns display regression results for the model presented in Equation (4) using parents' mean income and length of education as measures of schools' characteristics. In column 1 the dependent variable is compliance, i.e, enrollment in the assigned school where control units are artificially assigned nearby recipient schools. The other columns capture one of the margins of non-compliance from Equation (2), see the columns' titles. The models are estimated with origin-attendance-zone by year fixed effects and origin-attendance-zone-by-new-attendance-zone fixed effects. Standard errors are in parentheses and clustered on the level of origin-attendance-zones. The full output of the top panel can be found in Appendix Table B.4.



(a) Distance to school (b) Share in district school (c) Share in private school

Figure B.1: School distance and enrollment

The figures depict various statistics for distance and enrollment. Figure B.1a shows the cumulative distribution of distance to the district school and the nearest private school. Figures B.1b and B.1c plot the annual share of children enrolled respectively in the district school and in a private school. The sample consists of all children at age 7 between 2008 and 2015. For enrollment, the density measures are: low density, less than 1000 per sq. km; medium density, between 1000 and 5000 per sq. km, and; high density, more than 5000 per sq. km.

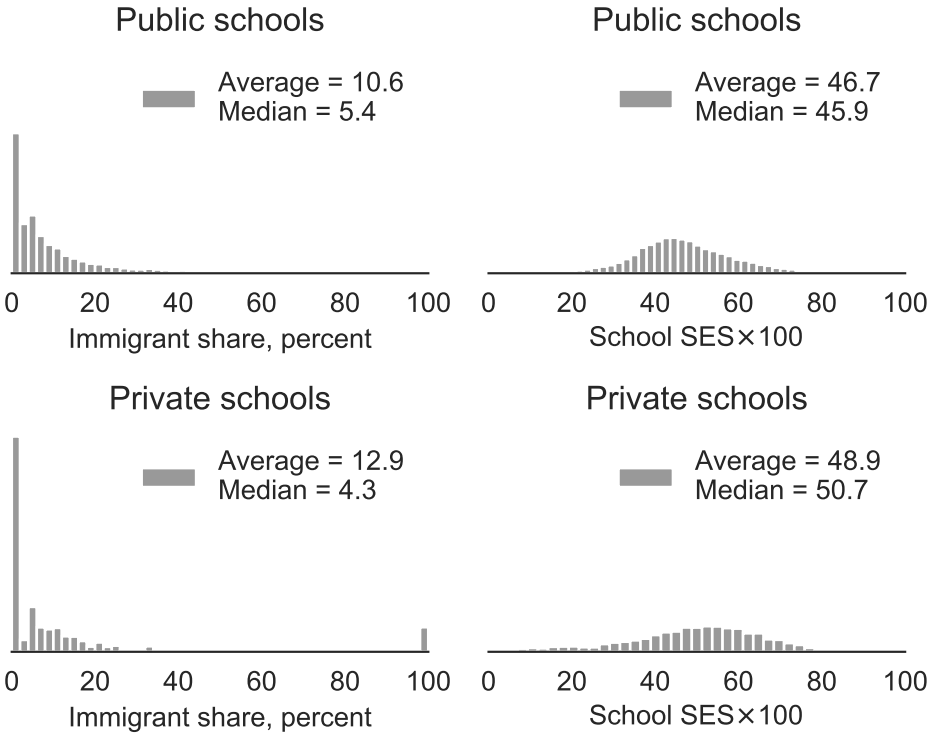


Figure B.2: Distribution of immigrant share and school-SES for private and public schools

The figures depict the distribution of school characteristics for public and private schools for the years 2008-2015. The mean immigrant share and school-SES are measured for seven-year old children within each school-year. The y-axis are standardized across plots.

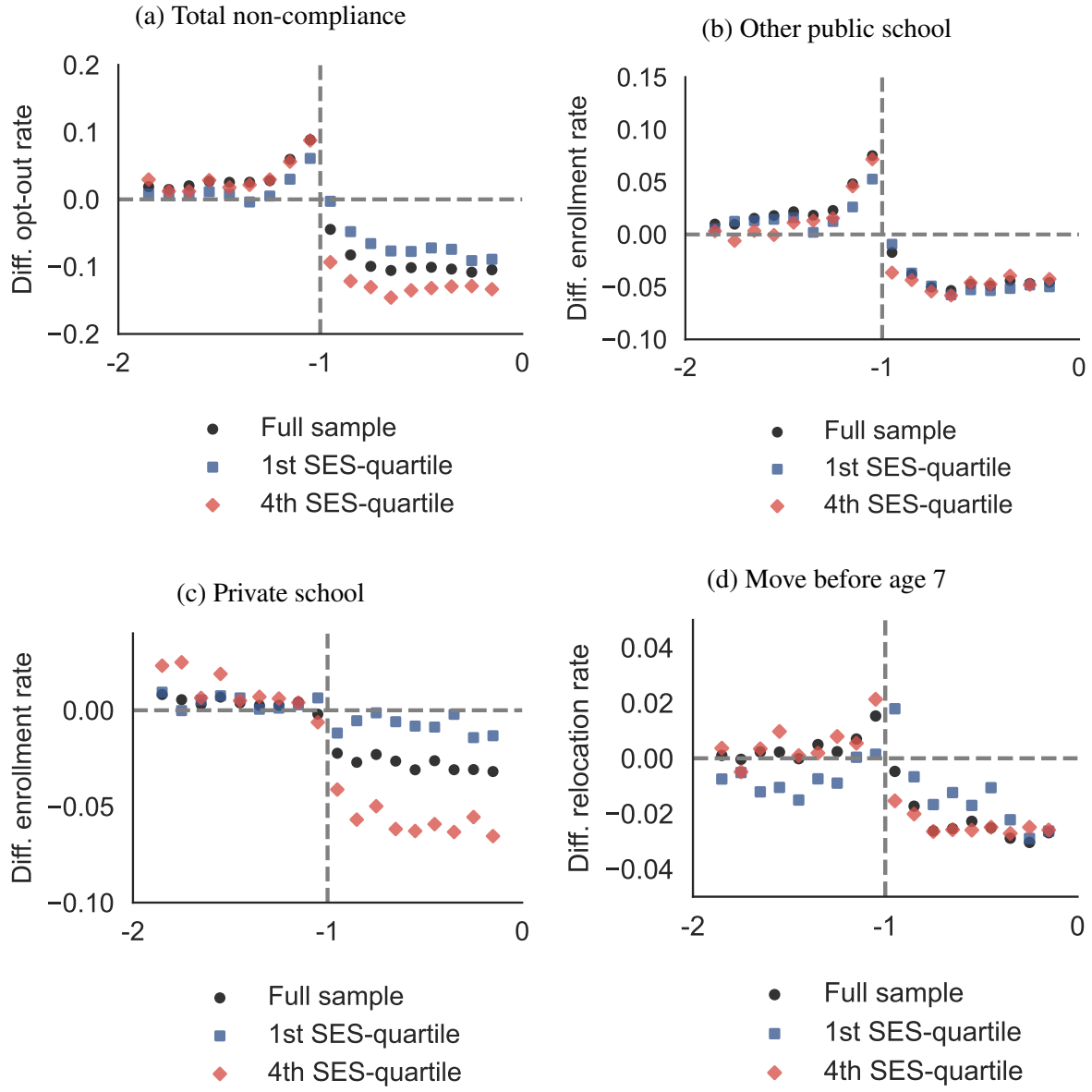
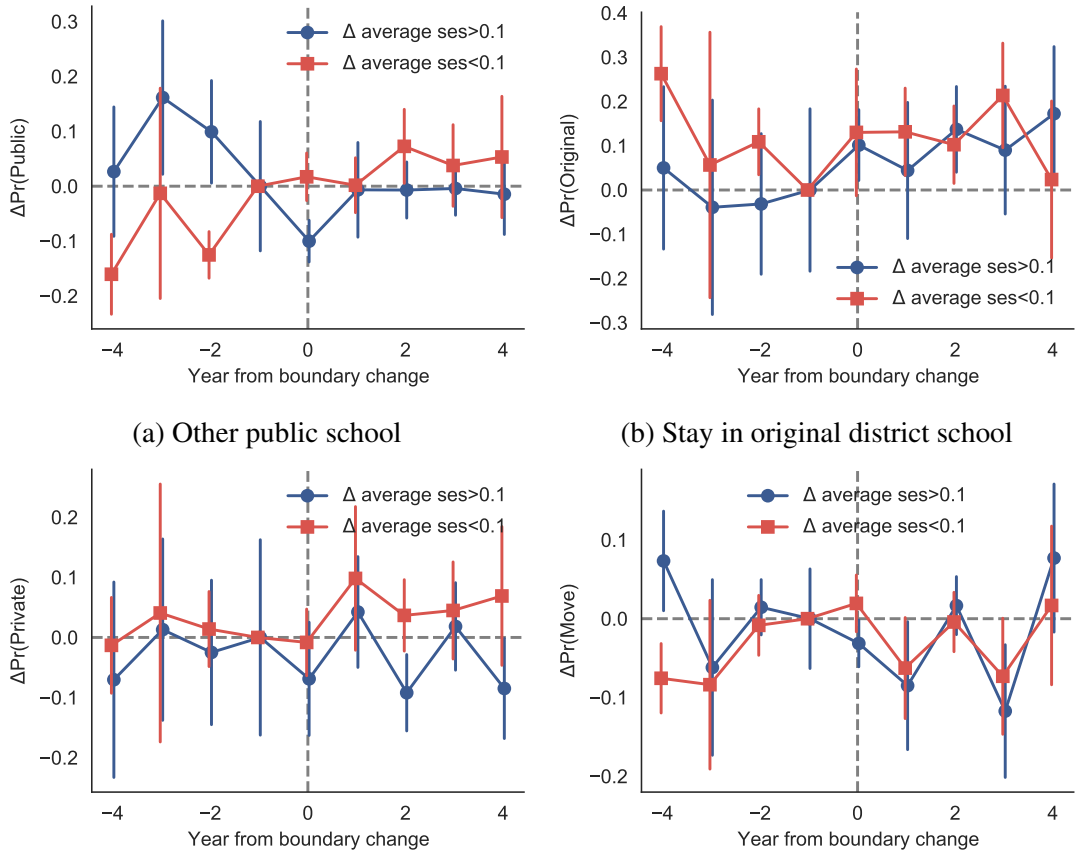


Figure B.3: Boundary difference in opt-out for low/high SES schools - by household SES

Each of the figures in the four panels depicts the boundary discontinuity estimates for the three margins of non-compliance in Equation (1) as well as the aggregate non-compliance. The dependent variables are displayed in the figure titles. A negative distance to the boundary signifies that the household is situated in the district of the two adjacent districts with the lower value of the school characteristic. The models are estimated with fixed effects at the boundary-year level. Results are centered on the left side of the boundary. The model is estimated separately for each quartile of household SES. The dependent variables in Figure B.3b, B.3c and B.3d are dummies that take the value of one if the observed child is respectively enrolled in another public school, enrolled in a private school, or moves out of the district between age 5 and 7. The mean difference is estimated in OLS and displayed in the lower-left corner of the figures.



(a) Other public school

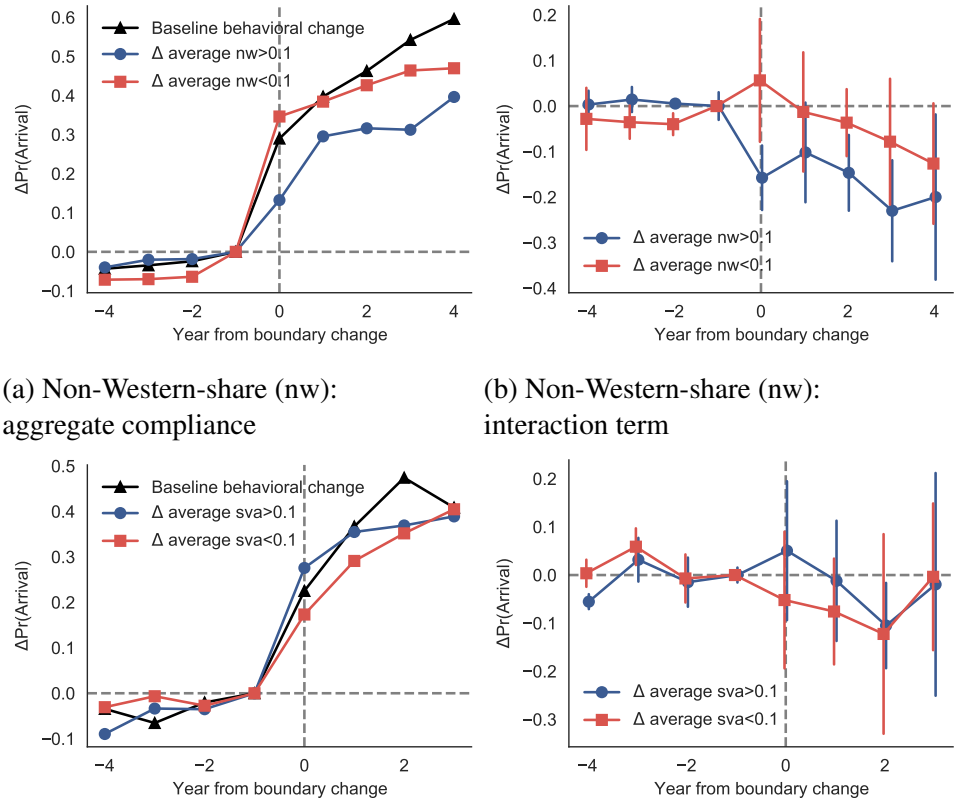
(b) Stay in original district school

(c) Private school enrollment

(d) Moving

Figure B.4: Response along enrollment margins by the change in school characteristic of assigned school

The figure display the interaction terms, β_τ and γ_τ from Equation (3), along with 95-percent confidence intervals. The parameters represent the difference in the likelihood of enrolling in the new district school when the average SES at a school level changes relative to the average arrival probability following a district change. The dependent variables of all figures are binary and measured at age 7 based on the district at age 7 for address at age 5. The models are estimated with origin-attendance-zone by year fixed effects and origin-attendance-zone-by-new-attendance-zone fixed effects. Standard errors are clustered on the level of municipalities. Results are centered on the year before the boundary change. Estimates from a simple before-after-DID are reported in the legends of Figure 5b.



(a) Non-Western-share (nw): aggregate compliance

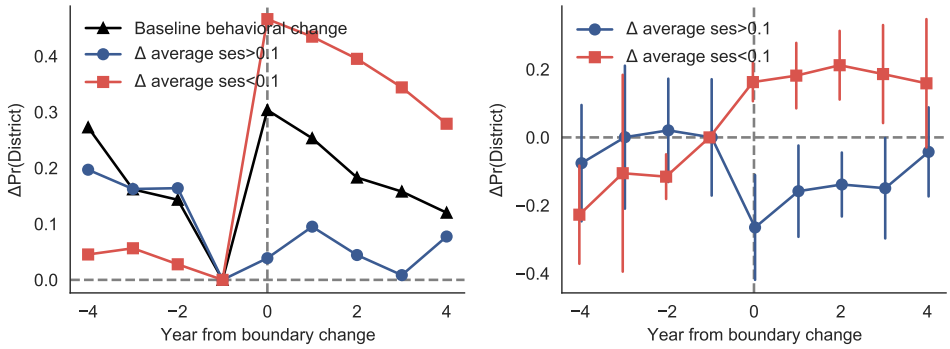
(b) Non-Western-share (nw): interaction term

(c) School value-added (sva): aggregate compliance

(d) School value-added (sva): interaction term

Figure B.5: Compliance as a function of changes to assigned school's non-Western share and school value-added

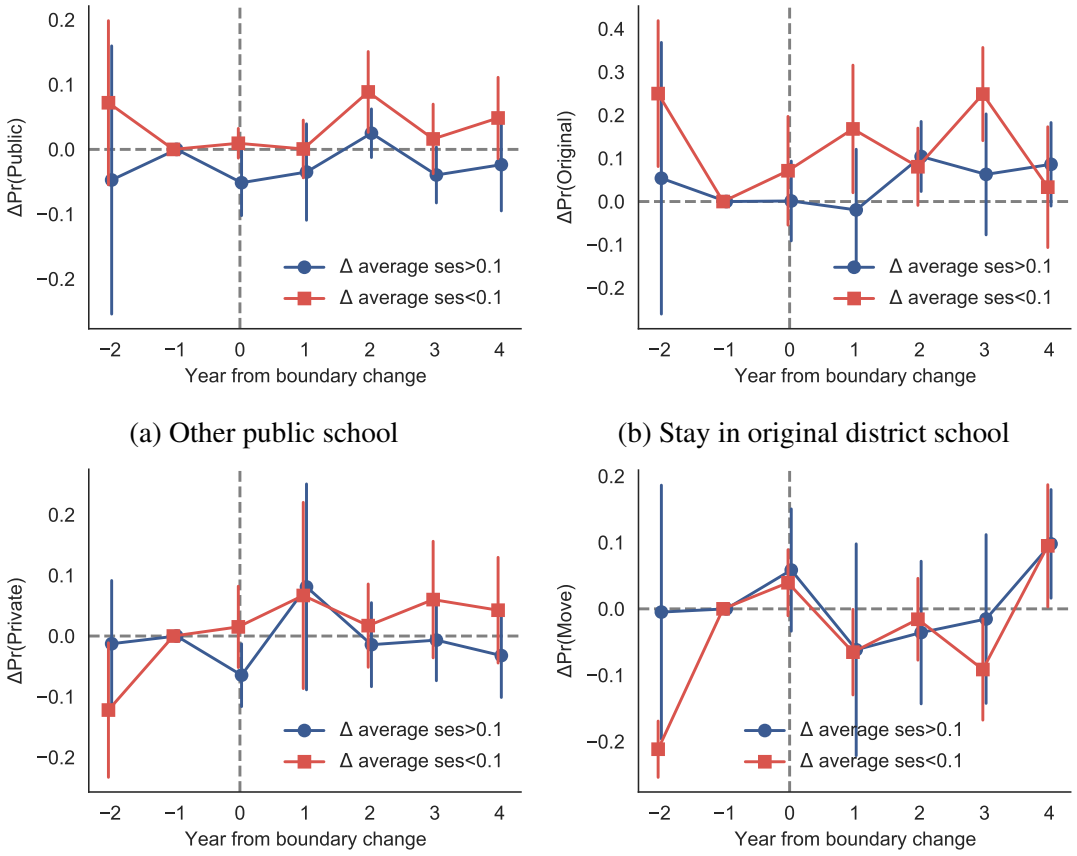
The figures on the left display changes in estimated compliance rates based on the model of Equation (3) estimated with different measures of school characteristics. The black lines depict the estimated α_τs from Equation (3), while the blue and red lines depict $\alpha_\tau + \gamma_\tau$ and $\alpha_\tau + \beta_\tau$ respectively. The figures to the right display the interaction terms, β_τ, and γ_τ, along with their 95-percent confidence intervals. The parameters represent the difference in the enrollment rate at the newly assigned school when the school characteristic at a school level change compared to the enrollment rate when school characteristics do not change. The dependent variable is binary and equals one if the child is enrolled in the district school at age 7 based on the district at age 7 for address at age 5. The y-axis denotes the excess probability of enrolling relative to baseline. The models are estimated with origin-attendance-zone by year fixed effects and origin-attendance-zone-by-new-attendance-zone fixed effects. Standard errors are clustered on the level of municipalities. Results are centered on the year before the district change. Estimates from a simple before-after-DID are reported in the legends of Figure 5b.



(a) Agg. non-compliance (b) Interaction parameters

Figure B.6: Alternative definition of opting out, School-SES

The figure display the interaction terms, β_τ and γ_τ, along with 95-percent confidence intervals. The parameters represent the difference in the likelihood of opting out of the designated school at age 7 when the average SES at a school level changes relative to the average arrival probability following a district change. The dependent variable is binary and measured at age 7 based on the district at age 7 for address at age 5. The models are estimated with origin-attendance-zone by year fixed effects and origin-attendance-zone-by-new-attendance-zone fixed effects. Results are centered on the year before the boundaries were redrawn. Estimates from a simple before-after-DID are reported in the legend.



(a) Other public school

(b) Stay in original district school

(c) Private school enrollment

(d) Moving

Figure B.7: Response along margins to change in the district by school characteristic when child aged 3 at time of boundary change

The figure display the interaction terms, β_{-}^{k} and β_{+}^{k}, along with 95-percent confidence intervals. The parameters represent the difference in the likelihood of enrolling in the new district school when the average SES at a school level changes relative to the average arrival probability following a district change. The dependent variables of all figures are binary and measured at age 7 based on the district at age 7 for address at age 3. The models are estimated with origin-attendance-zone by year fixed effects and origin-attendance-zone-by-new-attendance-zone fixed effects. Standard errors are clustered on the origin-attendance-zone level. Results are centered on the year before the district change.

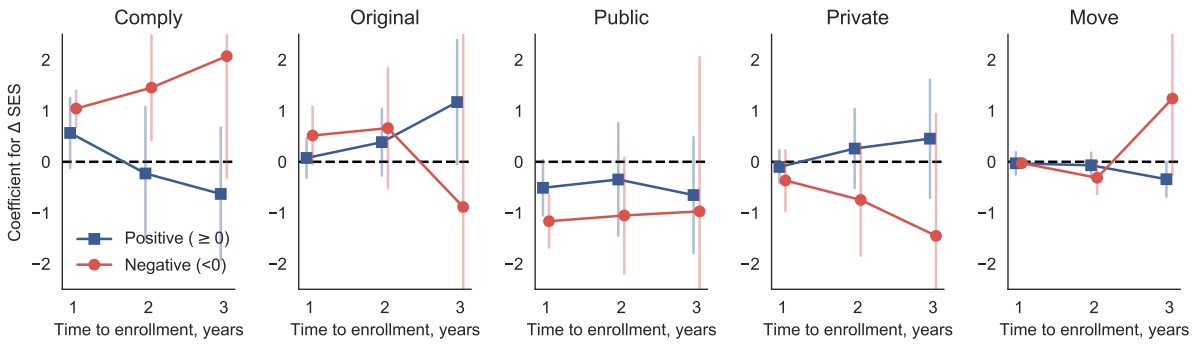


Figure B.8: DID-estimates as a function of time to enrollment

The figure displays how sensitive the results from estimating the model presented in Equation (4) are to changes in school-SES for children at younger ages. The younger the child, the longer time to exploit margins. The estimates at year 1, 2 and 3 correspond to a sampling of children of respectively 5, 4 and 3-year-olds with addresses in the original attendance zone. Thus, the estimates for the 5-year-olds are the same estimates as reported in the middle panel of Table 4. The models are estimated with origin-attendance-zone by year fixed effects and origin-attendance-zone-by-new-attendance-zone fixed effects. 95-percent confidence intervals are based on standard errors clustered at the level of municipalities.